

La prévision de la faillite fondée sur l'analyse financière de l'entreprise : un état des lieux

Catherine Refait

Citer ce document / Cite this document :

Refait Catherine. La prévision de la faillite fondée sur l'analyse financière de l'entreprise : un état des lieux. In: Économie & prévision, n°162, 2004-1. pp. 129-147;

doi : 10.3406/ecop.2004.6937

http://www.persee.fr/doc/ecop_0249-4744_2004_num_162_1_6937

Document généré le 17/06/2016

Résumé

La prédiction de la faillite des entreprises fait l'objet de nombreux travaux empiriques, depuis une trentaine d'années. Elle se fonde sur l'analyse économique et financière d'entreprises défaillantes et d'entreprises non défaillantes, afin de déterminer les variables, principalement comptables, qui distinguent au mieux les deux catégories de firmes. Nous proposons un état des lieux afin de rendre compte de l'efficacité relative des différentes méthodes de classification utilisées. Dans ce but, nous exposons la démarche commune tout en mettant en évidence les différentes modalités d'application empirique. Nous présentons le principe des techniques disponibles et une comparaison de leur performance, en mettant l'accent, de manière non exhaustive, à la fois sur les études fondatrices et sur les études les plus récentes. Mots-clés : prévision, faillite d'entreprise, analyse financière

Abstract

A Review of Business Failure Prediction Based on Financial Analysis of the Firm

Many studies of business failure prediction have been carried out in the last thirty years or so. Prediction is based on an economic and financial analysis of failing and non-failing firms with the aim of determining the variables, mainly of an accounting nature, that best distinguish the two categories. The purpose of this review is to assess the relative effectiveness of the different classification methods employed. To that end we identify the common elements while highlighting the different ways in which they are applied empirically. We describe the principle of the available techniques and compare their performance, focusing non-exhaustively on both the founding studies and on more recent research. Key-words : prediction, business failure, financial analysis

La prévision de la faillite fondée sur l'analyse financière de l'entreprise : un état des lieux

Catherine Refait^(*)

La prédiction de la faillite des entreprises fait l'objet de nombreux travaux empiriques, depuis une trentaine d'années. Elle se fonde sur l'analyse économique et financière d'entreprises défaillantes et d'entreprises non défaillantes, afin de déterminer les variables, principalement comptables, qui distinguent au mieux les deux catégories de firmes. Nous proposons un état des lieux afin de rendre compte de l'efficacité relative des différentes méthodes de classification utilisées. Dans ce but, nous exposons la démarche commune tout en mettant en évidence les différentes modalités d'application empirique. Nous présentons le principe des techniques disponibles et une comparaison de leur performance, en mettant l'accent, de manière non exhaustive, à la fois sur les études fondatrices et sur les études les plus récentes.

(*) Université Paris IX Dauphine – CERPEM CREFED.
E-mail : catherine.refait@dauphine.fr

Je tiens à remercier Mme. Bardos pour ses précieuses remarques. Merci également au rapporteur anonyme qui m'a permis d'améliorer la qualité de mon travail. Je demeure seule responsable du contenu de cet article.

L'accroissement actuel de la taille des entreprises en faillite, donc des montants de dettes concernés, rappelle vivement la nécessité de prévoir la défaillance. La protection des intérêts des créanciers, voire le rétablissement de la pérennité de l'entreprise, passent par l'anticipation des graves difficultés économiques et financières qu'une entreprise est susceptible de rencontrer. Actuellement le problème se pose aux banques avec une acuité particulière. Dans le cadre du Comité de Bâle et du ratio Mac Donough, les banques se devront en effet de proposer une évaluation systématique des risques qu'elles encourent. Ceci implique notamment une estimation précise de la probabilité de défaut de leurs clients entreprises, donc un éventuel remaniement de leurs méthodes d'évaluation (Altman, 2002).

Si l'analyse des causes de la faillite est plus ancienne, sa prédiction, c'est-à-dire l'évaluation quantitative du risque de défaillance, s'est principalement développée à partir de la fin des années soixante⁽¹⁾. L'approche la plus fréquente consiste à recourir à l'analyse financière afin de déterminer les variables, principalement comptables, qui différencient au mieux les entreprises défaillantes et les entreprises non défaillantes. L'objectif est d'établir une relation statistique stable entre les variables explicatives retenues et l'appartenance des entreprises à l'un des deux groupes. Différents outils sont à la disposition des auteurs, dont le plus fréquemment utilisé est l'analyse discriminante linéaire, mais les méthodes alternatives sont nombreuses.

L'objet de cet article est de présenter de manière synthétique les évolutions et la situation actuelle de la recherche, tant exploratoire qu'opérationnelle, en ce domaine. En proposant une explication succincte des différentes techniques utilisées et une comparaison des résultats obtenus, il souhaite offrir une vue d'ensemble des problèmes soulevés par la prédiction de la faillite. Il ne constitue cependant pas un recensement exhaustif des études réalisées. Non seulement un tel recensement eut été très difficile à réaliser compte tenu du nombre élevé de ces études, mais en plus il nous a semblé plus intéressant de mettre l'accent sur certains travaux représentatifs. Cet article s'inscrit donc dans la lignée de travaux tels que ceux de Malécot (1986), Dumontier (1991), voire Altman et Narayan (1997). Le développement récent de nouvelles méthodologies fondées sur l'intelligence artificielle et l'importance de leurs applications justifient qu'un nouvel état des lieux soit réalisé. Nous proposons à la fin de l'article un tableau récapitulatif des principales études citées, en privilégiant celles qui permettent une comparaison rigoureuse des différentes techniques disponibles. La première partie expose la méthodologie commune à l'ensemble des travaux considérés mais aussi les différentes modalités d'application empirique, dans la construction de l'échantillon, le choix des variables explicatives et la méthode de

validation des résultats. La deuxième partie met ensuite l'accent sur la variété des techniques de classification utilisées : les méthodes statistiques paramétriques et non-paramétriques ainsi que les méthodes d'intelligence artificielle. Nous concluons en proposant une synthèse de l'efficacité de chacune des techniques présentées.

Démarche commune et différentes modalités d'application empirique

Depuis les travaux de Tamari (1964), de nombreux auteurs ont tenté avec succès d'évaluer le risque de défaut des entreprises en se fondant sur leur analyse financière. Même si les techniques utilisées sont variées, le principe général qui sous-tend les diverses études est similaire : exploiter la connaissance *ex post* de l'avenir des entreprises. Les auteurs disposent des données comptables d'entreprises dont ils savent si elles ont été ou non défaillantes à la fin de la période d'observation. Ils parviennent ainsi à sélectionner les variables comptables les plus discriminantes puis à établir une relation statistique entre ces variables et l'état dichotomique d'être ou de ne pas être défaillant. Une probabilité de faillite est parfois également estimée. La qualité de l'indicateur de risque ainsi élaboré est jugée au regard des erreurs de classement : quel est le pourcentage d'entreprises défaillantes considérées comme saines – erreur de type I – et, inversement, quel est le pourcentage d'entreprises non défaillantes considérées comme risquées – erreur de type II ? Quatre étapes successives sont ainsi nécessaires à la prévision de la faillite : la construction d'un échantillon, c'est-à-dire la sélection d'entreprises, la sélection *a priori* des variables explicatives de la faillite, le choix du mode de classification des entreprises entre le groupe des défaillantes et celui des non-défaillantes que nous traiterons de façon détaillée dans la deuxième partie, enfin l'estimation de la qualité de la prévision effectuée. Nous explicitons ces étapes afin de mettre en évidence les distinctions entre les différentes études.

La construction de l'échantillon

L'élaboration de l'échantillon joue un rôle important dans la qualité de la prévision proposée. En premier lieu, la définition de la défaillance et le choix de l'horizon de prévision par les auteurs sont cruciaux, car ils déterminent l'objectif de l'indicateur construit ainsi que la sélection des entreprises qui constituent l'échantillon utilisé. Le choix des caractéristiques des entreprises sélectionnées est également primordial, puisqu'il conditionne l'homogénéité et la représentativité de l'échantillon et influence donc la qualité de la prévision.

Définition de la défaillance et choix de l' horizon de prévision

La définition de la défaillance est primordiale dans la construction de l'échantillon, puisqu'elle caractérise les deux populations qui fondent l'indicateur de risque. En effet, à partir d'un échantillon global d'entreprises, deux sous-échantillons doivent être construits : l'échantillon des entreprises dites "défaillantes" et l'échantillon des entreprises dites "non-défaillantes". La détermination de ce qu'est une entreprise défaillante est un problème délicat, susceptible de nuire à la robustesse de l'indicateur.

La majorité des auteurs considère la défaillance comme l'ouverture d'une procédure judiciaire⁽²⁾. Cependant, certains auteurs considèrent comme "défaillante" toute entreprise qui a connu un défaut de paiement, arguant que cet événement en lui-même préoccupe les créanciers. Le comité de Bâle met en effet l'accent sur l'estimation du risque de défaut et non sur celle du risque de faillite. La définition du défaut de paiement peut prendre plusieurs formes. Généralement, le risque de défaut réside en le non-respect par le débiteur de ses obligations financières : non-remboursement du capital ou non-versement des intérêts, violation d'un *covenant*⁽³⁾. Mais la définition peut être plus large et consister en une dégradation de la qualité de signature de l'entreprise : restructuration de la dette, diminution des dividendes versés, voire avis défavorable d'un audit ou encore détérioration du *rating* du débiteur.

Plusieurs possibilités sont alors envisageables. La plus traditionnelle consiste à opposer les entreprises qui ont fait l'objet d'une procédure collective aux autres entreprises. Il est également possible d'opposer les entreprises qui ont connu un défaut de paiement, y compris celles qui ont fait faillite, aux autres (Beaver, 1966). Au contraire, d'autres auteurs limitent l'échantillon des entreprises "défaillantes" à celles qui ont fait faillite. Ils excluent alors de la population des entreprises "non défaillantes" celles qui présentent une santé financière fragile – les entreprises "vulnérables", répondant à un critère plus large que le simple défaut de paiement – afin d'obtenir une discrimination plus marquée (Altman et Loris, 1976 ; Altman, 1977 ; Taffler 1982). La qualité de la prévision n'est cependant pas accrue de manière significative. Une dernière approche a été envisagée : prévoir l'ouverture d'une procédure judiciaire, mais en sélectionnant exclusivement pour les deux sous-échantillons des entreprises qui ont connu des défauts de paiement (Flagg *et alii*, 1991) ou plus généralement des entreprises vulnérables. La qualité de la prévision peut se trouver affaiblie⁽⁴⁾ à cause de la plus grande similitude des deux sous-échantillons⁽⁵⁾.

Parallèlement à la définition de la défaillance, le choix de l'horizon de prévision est important. Il

conditionne la date de la défaillance des entreprises et la date des données qui vont fonder la discrimination, l'écart de temps entre les deux étant l'horizon de prévision. Un arbitrage est à faire entre une échéance trop proche, dont l'intérêt est limité car elle n'autorise pas les décisions à même de limiter les pertes, voire d'éviter la faillite, et une échéance trop lointaine, qui interdit une prévision précise. Les études diffèrent légèrement sur ce point, mais très généralement, deux horizons sont choisis : un an et trois ans avant la défaillance.

Les caractéristiques des entreprises : représentativité et homogénéité de l'échantillon

La construction des deux sous-échantillons pose le double problème de la représentativité et de l'homogénéité des deux sous-échantillons. Afin que l'indicateur de risque puisse être appliqué à une entreprise quelconque, l'échantillon à partir duquel il est établi doit être représentatif de l'économie : du point de vue des secteurs d'activité, de la taille des entreprises mais également du rapport entre le nombre d'entreprises défaillantes et d'entreprises non-défaillantes. Cependant, cette représentativité crée une hétérogénéité qui est susceptible de créer un biais statistique : des facteurs explicatifs peuvent être masqués par des effets sectoriels ou des effets de taille. Afin de concilier ces deux exigences et d'améliorer la prédiction de la défaillance, plusieurs solutions ont été préconisées.

Soulignons en premier lieu que le problème ne se pose pas de manière identique selon la taille de l'échantillon. Si les premières études, de nature plus exploratoire, recouraient à un nombre restreint d'entreprises (66 pour Altman, 1968, par exemple), les études actuelles se fondent sur de larges échantillons (presque 1 000 entreprises pour Lennox, 1999, et plus de 40 000 pour le score de la Banque de France, cf. Bardos, 1998, par exemple). Lorsque le nombre d'entreprises exploitées est élevé, une bonne représentativité est possible. Elle doit cependant être vérifiée par la comparaison entre les caractéristiques de l'échantillon et celles de la population globale. Une autre manière d'obtenir un échantillon représentatif est de limiter la portée de l'indicateur à une population cible, définie par un secteur économique limité ou un intervalle restreint de taille des entreprises. Les entreprises de l'échantillon se situent alors dans ce secteur économique ou dans cet intervalle de taille (Trieschmann et Pinches, 1973 ; Altman, 1977 ; Altman et Loris, 1976 ; Calia et Ganuci, 1997). L'inconvénient d'une telle méthode est qu'elle exige que soit élaboré un indicateur par secteur et par taille. Se pose alors le problème du choix d'agrégation des secteurs économiques et de l'ampleur des intervalles de taille. Comment estimer le degré optimal d'homogénéité de l'échantillon ? Cependant, cette méthode présente l'avantage certain de concilier représentativité et homogénéité.

Si un échantillon hétérogène est utilisé, comment éviter le biais statistique qui en découle ? Une première possibilité est l'utilisation de "ratios relatifs". Rapporter la valeur prise par un ratio dans une entreprise donnée à sa valeur moyenne au sein du secteur considéré permet en effet d'atténuer la spécificité sectorielle (Lev, 1969 ; Izan, 1984, ou encore Platt et Platt, 1990). Platt et Platt (1991) élaborent deux indicateurs, à partir des mêmes données, mais l'un utilise les ratios relatifs, l'autre les ratios absolus. Le premier est plus performant (cf. tableau récapitulatif). Néanmoins, peut-être à cause de sa lourdeur, comme le supposent les auteurs, cette solution est peu utilisée. La méthode la plus usitée consiste à procéder par appariement, en faisant correspondre à toute entreprise défaillante une entreprise non-défaillante de même taille et appartenant au même secteur économique (Mossman *et alii*, 1998). Lorsqu'une contrainte relative à la disponibilité des données existe, cette méthode est celle qui réduit au mieux le biais statistique lié à l'hétérogénéité de la population.

L'inconvénient est que l'échantillon des entreprises défaillantes est alors nécessairement de la même taille que celui des entreprises non-défaillantes : un autre biais peut apparaître dans l'estimation des paramètres, car la structure de la population globale n'est pas respectée (Malécot, 1986). Ce problème se pose d'ailleurs quel que soit le mode d'échantillonnage à cause de la rareté de la défaillance. Or le résultat de certaines techniques de classification dépend de la taille respective des deux échantillons⁽⁶⁾. Afin de limiter ce biais, il est possible – sauf pour la méthode d'appariement – de rendre le rapport de taille des deux échantillons conforme à la probabilité de défaillance *a priori*. Cette probabilité correspond au taux de défaillance de la population globale ou de la population cible. Il est néanmoins plus fréquent⁽⁷⁾ que le biais créé soit corrigé lors de l'élaboration de l'indicateur, en tenant compte alors de la probabilité de défaillance *a priori*.

Le choix des ratios explicatifs de la défaillance

La prévision de la défaillance est réalisée à partir de données comptables, parfois des données de marché⁽⁸⁾. Une première sélection est réalisée *a priori*, afin de n'envisager que des variables considérées par les analystes financiers comme représentatives de la santé d'une entreprise (en général une cinquantaine). Puis, grâce aux techniques de classification statistique, une seconde sélection est opérée afin de ne conserver qu'un nombre réduit de variables, les plus discriminantes (entre 1 – Beaver, 1966 – et 20 – Calia et Ganugi, 1997).

Les variables explicatives retenues, majoritairement des ratios comptables, sont diverses. Rose et Giroux (1984) en ont recensé plus de 130 différentes. Cependant, les caractéristiques des entreprises prises en compte sont généralement similaires.

Conformément à l'enseignement de l'analyse financière, la rentabilité de l'entreprise (économique ou financière), la structure de son bilan et sa capacité de remboursement sont les trois éléments les plus corrélés à la défaillance.

La rentabilité économique met en relation une variable de résultat économique avec l'actif total, le capital engagé⁽⁹⁾, les immobilisations productives ou le capital économique⁽¹⁰⁾. Le résultat considéré peut être le résultat d'exploitation (Altman, 1968 ; Taffler, 1982 ; Flagg *et alii*, 1991 ; Michalopoulos *et alii*, 1993), le résultat global – pour la mesure du *return on assets* par exemple (Weiss, 1996) – ou le résultat net (Burgstahler *et alii*, 1989 ; Calia et Ganugi, 1997). La rentabilité financière met en rapport une variable de résultat avec le capital financier. Elle est souvent exprimée par le ratio résultat global sur capitalisation boursière (*return on equity*) (Mensah, 1984).

La structure du bilan met en rapport un élément d'actif et un élément du passif ou deux éléments de passif, voire deux éléments d'actif. Elle permet de rendre compte d'éventuels déséquilibres qui peuvent mener l'entreprise au non-respect de ses engagements financiers. Elle est principalement appréhendée de trois manières. Le rapport entre l'actif total et l'endettement total rend compte de la solvabilité de l'entreprise (Deakin, 1972 ; Rose et Giroux, 1984 ; Burgstahler *et alii*, 1989 ; Michalopoulos *et alii*, 1993 ; Altman *et alii*, 1994) ; la solvabilité est également mesurée par le ratio endettement total sur fonds propres (Mensah, 1984 ; Weiss, 1996 ; Flagg *et alii*, 1991 ; Platt et Platt, 1991). Le rapport entre les actifs liquides – souvent hors stocks – et le passif exigible à court terme mesure le risque d'illiquidité (Deakin, 1972 ; Houghton, 1984 ; Burgstahler *et alii*, 1989 ; Michalopoulos *et alii*, 1993). Enfin, le ratio dettes à court terme sur dettes à long terme permet d'intégrer la nature de l'endettement (Platt et Platt, 1991). D'autres ratios de structure de bilan sont envisagés, plus marginalement, comme le ratio dettes à long terme sur fonds propres (Houghton, 1984 ; Rose et Giroux 1984) ou le ratio actifs liquides sur actif total (Deakin, 1972 ; Taffler, 1982).

Le dernier des trois facteurs les plus explicatifs, la capacité de remboursement, est mesuré de différentes manières : par le ratio *cash flow*⁽¹¹⁾ sur dette totale (Deakin, 1972 ; Frydman *et alii*, 1985 ; Calia et Ganugi, 1997), par le ratio intérêts dus sur chiffre d'affaires (Rose et Giroux, 1984), par le ratio revenu d'exploitation sur intérêts dus (Calia et Ganugi, 1997)⁽¹²⁾.

Parallèlement à l'utilisation standard de ratios comptables, certains auteurs ont tenté d'exploiter des données différentes. Une première idée a été de prévoir la faillite à partir des flux de trésorerie de l'entreprise, donc de grandeurs absolues et non plus

de grandeurs relatives (Gentry *et alii*, 1985 ; Aziz *et alii*, 1988 ; Reilly, 1990). La faillite est en effet supposée survenir lorsque la valeur de l'entreprise devient trop faible pour qu'un créancier accepte de continuer à la financer en cas de défaut de paiement. Le recours aux flux de trésorerie est donc légitime puisque la valeur d'une entreprise est égale à la somme actualisée de ses *cash flows* futurs. Ajouter des flux de trésorerie aux ratios comptables améliore la qualité de la prévision de la faillite, quelle que soit la technique de classification utilisée, dans les trois modèles cités.

Une autre idée a été de prendre en compte l'évolution du cours de l'action de l'entreprise (Beaver, 1966 ; Altman et Brenner, 1981 ; Clark et Weistein, 1983). Le cours se détériore de un an à trois ans avant le dépôt de bilan. Aharony *et alii* (1980) ont élaboré un modèle de prédiction de la faillite en se fondant sur la variation du taux de rendement des actions. Ils constatent que plus la faillite est imminente plus le risque spécifique⁽¹³⁾ des titres s'accroît. Les deux méthodes ont été utilisées par Mossman *et alii* (1998) : elles procurent une prévision de la faillite moins performante qu'une analyse discriminante menée sur les données comptables des mêmes entreprises. Le recours aux données de marché est peu courant, ne serait-ce que parce qu'il exclut la majorité des PME, notamment en France.

Enfin, conjointement aux données comptables, des informations liées à l'organisation de l'entreprise ou à la nature de son financement, par exemple, peuvent être utilisées. Citons l'étude de Foglia *et alii* (1998) qui introduit comme variable explicative d'une analyse discriminante le nombre de banques auprès desquelles les entreprises sont endettées. Le lien positif qui existe entre cette variable et le risque de l'entreprise permet d'améliorer la prévision de la défaillance⁽¹⁴⁾.

La méthode de validation

Une fois l'échantillon constitué, il convient d'établir un lien statistique entre l'appartenance à l'un des deux groupes et les variables explicatives retenues. Différentes méthodes de classification sont alors disponibles. Leur étude fait l'objet de la partie suivante. La qualité de prévision de l'indicateur de risque ainsi élaboré doit être mesurée. Un dernier élément de divergence entre les différentes études apparaît alors : la méthode de validation.

L'ensemble des études juge cependant l'efficacité de leur indicateur par une mesure commune : le taux de bons classements qu'il procure. Le taux de bons classements des entreprises non-défaillantes est la part des entreprises non défaillantes correctement classées dans l'ensemble des entreprises non-défaillantes considérées. Le taux de bons classements des entreprises défaillantes est la part des entreprises défaillantes correctement classées dans l'ensemble des entreprises défaillantes

considérées. Une mesure globale est également envisageable : la moyenne des taux de bons classements pondérés par l'effectif respectif de chaque sous-échantillon (taux de bons classements apparent) ou pondérés par les probabilités *a priori* de défaillance et de non-défaillance.

La divergence entre les études se situe dans la construction de l'échantillon qui sert à calculer les taux de bons classements, nommé "échantillon test". Les auteurs qui disposent d'une très vaste base de données mesurent les taux de bons classements à partir d'un échantillon d'entreprises défaillantes et d'entreprises non-défaillantes distinct de l'échantillon qui a servi à l'élaboration de l'indicateur (l'échantillon initial) : Bardos (1998a) et Varetto (1998) par exemple.

Dans d'autres études, l'échantillon qui sert à la validation est l'échantillon initial (Platt et Platt, 1991 ; Altman *et alii*, 1994). Une telle validation est sujette à caution : l'indicateur élaboré peut être performant pour l'échantillon initial et ne pas l'être pour l'ensemble des entreprises ; le taux de bons classements ainsi calculé – le taux apparent – ne constitue pas un estimateur sans biais du taux de bons classements de la population globale. Or c'est en fait ce taux – le taux réel de bons classements – et uniquement ce taux qui rend pleinement compte de la qualité d'un indicateur du risque de faillite. Aussi, la majorité des auteurs optent-ils pour des méthodes de validation qui permettent de réduire le biais dans l'estimation du taux réel de bons classements lorsqu'ils ne disposent pas d'un échantillon test distinct de l'échantillon initial : validation croisée, *Jack-knife* et *Bootstrap*⁽¹⁵⁾.

Les trois méthodes se fondent sur un rééchantillonnage des n entreprises qui constituent l'échantillon initial. La validation croisée d'ordre V consiste à découper l'échantillon initial en V sous-échantillons de taille approximativement égale, puis à estimer l'indicateur de risque à partir de $V-1$ groupes pour enfin calculer le taux de bons classements sur le $V^{\text{ième}}$ groupe. L'opération est réalisée V fois, chaque sous-groupe servant successivement au calcul d'un taux de bons classements. Le taux de bons classements théorique est estimé par la moyenne des V taux de bons classements empiriques. Le *Jack-knife* consiste à construire n sous-échantillons de $n-1$ entreprises, en négligeant successivement chacune des n entreprises de l'échantillon initial. À partir des n sous-échantillons sont estimés n indicateurs de risque pour lesquels n taux de bons classements sont calculés sur l'échantillon initial. Le taux de bons classements théorique est estimé par la moyenne des n taux de bons classements empiriques. Le *bootstrap* consiste à construire un échantillon à partir d'un tirage aléatoire avec remise des entreprises de l'échantillon initial. L'échantillon construit est généralement de même taille que l'échantillon

initial. Un indicateur est élaboré à partir de l'échantillon construit et le taux d'erreur de classement qu'il procure est calculé sur l'échantillon initial. Un taux d'erreur apparent est également calculé sur l'échantillon construit. L'écart entre les deux taux d'erreur de classement est mesuré. L'opération est répétée un nombre de fois suffisant pour réduire voire éliminer le biais. Le taux d'erreur de classement théorique est estimé par le taux d'erreur apparent calculé sur l'échantillon initial auquel s'ajoute la moyenne empirique des écarts.

Il convient donc de comparer les performances des différentes études avec circonspection, le niveau des taux de bons classements dépendant du mode de validation choisi par les auteurs. Cependant, les trois méthodes de rééchantillonnage citées permettent non seulement de calculer une moyenne des taux de bons classements mais également un intervalle de confiance de cette moyenne. La précision de la prévision est donc connue et peut être comparée d'une étude à l'autre.

Les différentes techniques de classification utilisées pour la prévision de la faillite

L'objet commun des modèles de prévision de la faillite est de tenter, grâce aux ratios comptables sélectionnés, d'affecter une entreprise quelconque à l'un des deux groupes : les entreprises défaillantes et les entreprises non-défaillantes. Les études peuvent être répertoriées en trois classes, selon la méthode de classification utilisée. Les plus nombreuses recourent aux méthodes de classification statistiques paramétriques : les analyses discriminantes – qu'elles soient linéaires (Altman, 1968) ou quadratiques (Lachenbruch *et alii*, 1973) – et les techniques économétriques sur données qualitatives (Ohlson, 1980 ; Zmijewski, 1984). D'autres études recourent à des méthodes statistiques non-paramétriques comme le partitionnement récursif (Frydman *et alii*, 1985) et les estimateurs à noyau (Calia et Ganuci, 1997). Plus récemment ont été développées des analyses s'inspirant de l'intelligence artificielle comme les réseaux de neurones (Altman *et alii*, 1994 ; Bardos et Zhu, 1997) ou les algorithmes génétiques (Varetto, 1998)⁽¹⁶⁾. Nous exposons les différentes techniques à la lecture des articles fondateurs puis nous présentons les diverses améliorations ultérieures.

Les études fondées sur les méthodes paramétriques de classification statistique

Les méthodes paramétriques de classification statistique établissent une relation fonctionnelle entre les variables explicatives – dont la loi de distribution est supposée connue – et la variable

expliquée, relation dont la forme est donnée *a priori*. Trois grandes familles de méthodes existent : la méthodologie unidimensionnelle fondatrice de Beaver (1966), l'analyse discriminante (linéaire et non linéaire) et les régressions sur variables qualitatives.

La méthodologie unidimensionnelle de Beaver, 1966

Bien que les premiers travaux relatifs à la prédiction des faillites d'entreprises à partir de données comptables soient l'œuvre de Tamari (1964), les articles de Beaver (1966) puis d'Altman (1968) semblent avoir été le réel point de départ et la référence des nombreuses études empiriques publiées depuis trente ans. Beaver (1966) élabore une classification dichotomique unidimensionnelle, c'est-à-dire fondée sur un ratio unique. Sa méthode, de nature exploratoire, n'est plus appliquée actuellement. Nous la présentons donc succinctement.

L'objectif de Beaver est de classer les entreprises sur la base du ratio le plus discriminant. Pour ce faire, il sélectionne initialement pour chaque entreprise de son échantillon différents ratios comptables, censés, d'après l'analyse financière, être d'autant plus élevés que la santé financière des entreprises est saine. Il estime la valeur prédictive de chaque ratio de la manière suivante. Dans un premier sous-échantillon, il classe les entreprises en fonction de la valeur prise par un ratio particulier. Il choisit un seuil critique : toute entreprise présentant un ratio inférieur au seuil est considérée comme défaillante et au contraire toute entreprise présentant un ratio supérieur est considérée comme non-défaillante. Le seuil critique est déterminé de manière à maximiser le taux de bons classements dans le premier sous-échantillon. Un classement des entreprises du second sous-échantillon est ensuite réalisé à partir du seuil critique précédemment déterminé et un autre taux de bons classements est calculé. C'est ce taux qui fonde la sélection finale du ratio le plus discriminant.

À un an de la défaillance, le ratio *cash flow* sur endettement total est le plus discriminant ; il permet de classer correctement 87% des entreprises. Ce ratio est également celui qui fournit la meilleure prévision trois ans avant la faillite : le taux de bons classements est alors égal à 77%. La méthode de Beaver (1966) fournit donc un indicateur à la fois simple et efficace. Le manque de robustesse lié à l'unicité du ratio utilisé explique sans doute que cette méthode n'ait été que rarement exploitée par la suite (Deakin, 1972 ; Gebhardt, 1980). Les analyses multidimensionnelles, permettant une description plus riche de la situation de l'entreprise, sont maintenant utilisées de manière systématique.

Dès 1968, Altman affirme qu'une analyse unidimensionnelle rend mal compte de la complexité du processus de défaillance. Aussi est-il le premier à exploiter simultanément plusieurs ratios, au travers d'une analyse discriminante linéaire multidimensionnelle⁽¹⁷⁾. Afin de classer une entreprise parmi l'un des deux groupes d'entreprises : défaillantes ou non-défaillantes, Altman (1968) utilise la règle de décision simple qui consiste à affecter l'entreprise au groupe dont elle est le plus proche. Il recourt à l'analyse discriminante de Fisher fondée sur un critère métrique. L'analyse aboutit à la construction d'une fonction appelée score, combinaison linéaire des variables explicatives retenues, dont la réalisation exprime le niveau de risque de l'entreprise.

Des ratios, au nombre de K , sont sélectionnés initialement au sein de l'ensemble vaste des ratios susceptibles d'être explicatifs de la défaillance. Sont conservés ceux qui limitent le problème de multicollinéarité tout en possédant le pouvoir de discrimination le plus élevé. L'ensemble des entreprises peut ainsi être représenté dans l'espace \mathcal{R}^K . L'objectif de l'analyse discriminante est de séparer l'espace \mathcal{R}^K en deux sous-espaces : le sous-espace des entreprises défaillantes et le sous-espace des entreprises non-défaillantes. Chaque sous-espace contient des entreprises de l'autre groupe, puisque les K ratios ne caractérisent qu'imparfaitement les deux catégories d'entreprises. L'analyse discriminante vise à déterminer l'hyperplan H^* qui sépare au mieux les deux groupes. H^* doit donc être tel qu'à la fois le nombre d'entreprises défaillantes présentes dans le sous-espace des entreprises non-défaillantes soit minimal et que le nombre d'entreprises non-défaillantes incluses dans le sous-espace des entreprises défaillantes soit minimal. L'entreprise dont on désire connaître le risque de faillite – que nous appellerons l'entreprise A – est ensuite affectée au groupe qui correspond au sous-espace où ses coordonnées x_A la situent. Cette démarche revient à rattacher l'entreprise A au groupe dont le centre lui est le plus proche. Il convient donc d'établir une mesure de distance : la distance à un groupe est définie comme la distance au point moyen du groupe, suivant une métrique particulière⁽¹⁸⁾.

Soit $S(x)=0$ l'équation de l'hyperplan H^* dans l'espace \mathcal{R}^K . S est une combinaison linéaire des K ratios considérés. La valeur prise par la fonction $S(\cdot)$ au point x_A détermine l'affectation de l'entreprise à partir du vecteur x_A de ses K ratios : si $S(x_A)>0$, l'entreprise A est classée parmi les non-défaillantes, si $S(x_A)<0$, l'entreprise A est classée parmi les défaillantes et si $S(x_A)=0$ il est impossible de classer l'entreprise A . La valeur prise par la fonction score fournit également un indicateur du risque de faillite de l'entreprise. L'analyse discriminante

linéaire permet ainsi de raisonner dans un espace à une dimension au lieu de K , nombre de variables explicatives exploitées. Elle permet également de connaître directement la contribution de chaque variable au risque de l'entreprise.

L'analyse d'Altman (1968) se révèle plus efficace que celle menée par Beaver (1966) un an avant la défaillance puisqu'elle aboutit à un taux de bons classements global de 95% sur l'échantillon initial. Cependant, calculé sur un échantillon test distinct de l'échantillon initial, le taux de bons classements un an avant la défaillance est égal à 82% et donc inférieur au taux fourni par l'analyse de Beaver (qui était égal à 87%, calculé sur un échantillon test distinct de l'échantillon initial). Il peut sembler surprenant qu'ajouter des variables explicatives n'améliore pas la prévision. Cependant, comparer l'efficacité de deux prévisions réalisées à partir de deux échantillons différents est délicat, ce d'autant plus que leur taille est réduite (158 entreprises pour l'étude de Beaver, 1966, et 66 pour celle d'Altman, 1968). De plus leur objectif est différent : Beaver (1966) désire prévoir tout défaut de paiement alors qu'Altman (1968) se concentre sur le dépôt de bilan. Une comparaison entre les deux études ne peut donc être faite qu'avec prudence. Deakin (1972), en confrontant les deux méthodes à partir d'un même échantillon, tend à montrer la supériorité de l'analyse discriminante linéaire pour prévoir le dépôt de bilan. À un an de la faillite, elle classe correctement 87% des entreprises, alors que la classification unidimensionnelle élaborée par Beaver ne classe correctement que 80% des entreprises.

Plusieurs inconvénients sont liés à l'utilisation de l'analyse discriminante linéaire. Des modifications importantes ont été réalisées pour y remédier et elle a été la technique la plus usitée pendant 30 ans⁽¹⁹⁾, tous pays confondus⁽²⁰⁾. La première modification de la méthode est liée à l'intégration des coûts d'erreurs de classement et de la probabilité de défaut *a priori*. En effet, il peut tout d'abord être important pour le décideur – principalement un banquier ou un assureur crédit – de prendre en compte le fait que le coût des erreurs de type I (considérer qu'une entreprise défaillante est non-défaillante) est différent du coût des erreurs de type II (considérer qu'une entreprise non-défaillante est défaillante). Le coût d'une erreur de type I – noté c_D^{ND} – représente la perte totale ou partielle sur les nouveaux crédits octroyés à cause d'une mauvaise estimation du risque. Le coût d'une erreur de type II – noté c_{ND}^D – représente un coût d'opportunité pour la banque⁽²¹⁾. Par ailleurs, le seuil de décision présenté ci-dessus ne tient pas compte de la probabilité *a priori* qu'une entreprise soit défaillante ou non-défaillante⁽²²⁾. Si, par exemple, les deux types d'erreur ont le même coût, mais que la probabilité de défaut *a priori* est faible, une entreprise située à équidistance des deux

groupes devrait être considérée comme non-défaillante.

La règle de décision de Bayes permet d'utiliser la fonction score tout en prenant en compte les coûts d'erreur de classement et la probabilité de défaut *a priori*. Notons π_j la probabilité *a priori* qu'une entreprise quelconque appartienne au groupe j , où j peut désigner le groupe des défaillantes ($j=D$) ou celui des non-défaillantes ($j=ND$). Notons $f(x_A | j)$ la fonction de densité du vecteur des coordonnées x_A de l'entreprise A conditionnelle à l'appartenance au groupe j . L'expression $f(x_A | j)$ correspond à la probabilité conditionnelle que l'entreprise soit caractérisée par les coordonnées x_A alors qu'elle appartient au groupe j . Alors $f(j|x_A)$, la probabilité *a posteriori* pour l'entreprise A d'appartenir au groupe j sachant x_A , se déduit du théorème de Bayes :

$$f(j|x_A) = \frac{f(x_A | j)\pi_j}{f(x_A)}$$

Les probabilités *a posteriori* ne donnent pas lieu directement à une classification binaire des entreprises mais à la construction de classes de risques⁽²³⁾.

Si les coûts d'erreur de classement sont intégrés au critère de décision bayésienne, A est affectée au groupe des défaillantes si et seulement si : $f(x_A | D)c_D^{ND}\pi_D \geq c_{ND}^D\pi_{ND}f(x_A | ND)$. Cela signifie que la proportion des entreprises affectées au groupe des entreprises défaillantes croît avec le coût associé au mauvais classement d'une entreprise défaillante. Au contraire, plus le fait de considérer une entreprise non-défaillante comme défaillante est coûteux, plus la proportion d'entreprises affectées au groupe des non-défaillantes augmente.

Si les fonctions de densité des deux sous-populations sont multinormales et homoscedastiques, la fonction score $S(\cdot)$ permet de classer une entreprise A à partir du vecteur x_A de ses K ratios en fonction de la règle bayésienne de la manière suivante :

$$\text{si } S(x_A) > \ln \left(\frac{c_D^{ND}\pi_D}{c_{ND}^D\pi_{ND}} \right)$$

A est classée parmi les non-défaillantes,

$$\text{si } S(x_A) < \ln \left(\frac{c_D^{ND}\pi_D}{c_{ND}^D\pi_{ND}} \right)$$

A est classée parmi les défaillantes.

Une lecture directe de la fonction score permet ainsi la classification des entreprises. Soulignons cependant que la détermination d'un seuil de classification est secondaire par rapport à la valeur de la fonction score en-elle même qui fournit une indication du risque de l'entreprise. La première raison en est la difficulté d'évaluer les coûts d'erreur

de classement. La seconde raison est qu'une fonction score a pour fonction d'aider à la décision et non de fonder seule la décision.

D'autres modifications ont été apportées, afin de limiter les inconvénients liés à la linéarité de la fonction score construite par l'analyse discriminante de Fisher. Cette caractéristique implique que chaque ratio influence toujours dans le même sens le risque de faillite. Or certaines variables n'ont pas un impact monotone sur la santé de l'entreprise. Par exemple, le lien entre la taille de l'entreprise et son risque est croissant pour les petites entreprises puis il est décroissant au delà d'une certaine taille⁽²⁴⁾. De plus, la linéarité signifie que l'élasticité du score par rapport au ratio est monotone. Le lien entre le risque et les ratios choisis est donc soumis à des contraintes structurelles qui peuvent nuire à l'efficacité de l'analyse discriminante. Pour lever ces contraintes, Bardos (1989) applique à la prévision de la faillite la méthode DISQUAL, établie par Saporta (1977). Les variables comptables continues sont réparties en intervalle et ainsi transformées en variables qualitatives. L'absence de monotonie dans le lien entre le niveau des ratios et le niveau de risque de l'entreprise peut alors être pris en compte. L'indicateur ainsi construit est cependant moins performant que le score construit par une discrimination linéaire appliquée aux mêmes données. La plus grande robustesse du score est invoquée par Bardos (1989) pour expliquer ce résultat.

L'analyse discriminante linéaire présente un autre inconvénient majeur : elle requiert des conditions statistiques strictes. Les variables comptables utilisées doivent en effet suivre une loi multinormale et leurs matrices de variance-covariance doivent être les mêmes pour l'échantillon des entreprises défaillantes et pour celui des entreprises non-défaillantes⁽²⁵⁾. Face à la contrainte d'homoscédasticité, certains auteurs ont recouru aux analyses discriminantes quadratiques, qui n'exigent que l'hypothèse de multinormalité des ratios (Lachenbruch *et alii*, 1973 ; Marks et Dunn, 1974 ; Rose et Giroux, 1984, par exemple). À notre connaissance, l'analyse discriminante quadratique n'est plus utilisée depuis le milieu des années quatre-vingt. Si théoriquement elle est plus pertinente que l'analyse linéaire en cas d'hétéroscédasticité, dans les faits elle est moins performante, principalement pour deux raisons. Premièrement, l'absence de multinormalité des ratios est beaucoup plus nuisible à l'efficacité de l'analyse quadratique qu'à celle de l'analyse linéaire (Lachenbruch, 1975) ; deuxièmement, même en cas de respect de l'hypothèse de multinormalité, l'analyse discriminante quadratique n'est performante que si elle est appliquée à un échantillon très large. La complexité du modèle quadratique rend par ailleurs difficile l'analyse du rôle joué dans

le processus de défaillance par les différentes variables considérées.

L'analyse discriminante linéaire reste actuellement la méthode la plus utilisée. La fonction score, grâce à l'analyse de la contribution de chaque ratio, constitue une aide à la compréhension de la faillite. De surcroît, grâce à sa robustesse temporelle, elle offre la meilleure classification. De nombreuses études – principalement dans les années quatre-vingt dix – confrontent l'analyse discriminante linéaire aux autres techniques de classification (Collins, 1980 ; Rose et Giroux, 1984 ; Altman *et alii*, 1994 ; Calia et Ganuci, 1997 ; Kira *et alii*, 1997 ; Bardos et Zhu, 1997 ; Varetto, 1998). La majorité aboutit au résultat suivant : les taux de bons classements obtenus par l'analyse discriminante linéaire sont supérieurs à ceux fournis par les autres méthodes pour les mêmes données. La supériorité de la première n'est cependant pas systématique (Bardos et Zhu, 1997 ; Calia et Ganuci, 1997, notamment).

Les modèles de prédiction recourant aux régressions sur variables qualitatives

Face à la contrainte de multinormalité – rarement respectée empiriquement – des analyses discriminantes linéaires et quadratiques, certains auteurs ont préféré recourir à d'autres méthodes. La première possibilité est l'utilisation d'autres techniques paramétriques, qui supposent une distribution différente des variables comptables : les techniques économétriques sur variables qualitatives que sont le modèle *Probit* et le modèle *Logit*. Ils diffèrent dans la distribution supposée des erreurs : une loi logistique pour le *Logit* et une loi normale pour le *Probit*. La régression *Logit* a été appliquée à la prévision de la faillite par Ohlson (1980), puis notamment par Mensah (1984), Aziz *et alii* (1988), Bardos (1989), Burgstahler *et alii* (1989), Flagg *et alii* (1991), Platt et Platt (1991), Weiss (1996), Bardos et Zhu (1997), Mossman *et alii* (1998). Le modèle *Probit* a été plus rarement utilisé (Zmijewski, 1984 ; Gentry *et alii*, 1985 ; Lennox, 1999).

Dans les deux modèles, la variable endogène y est une variable qualitative, ici dichotomique⁽²⁶⁾ : elle prend la valeur 0 ou 1, selon que l'entreprise est défaillante ou non. Supposons que y soit nul pour les entreprises défaillantes et égal à un pour les entreprises non-défaillantes. Le vecteur x des variables exogènes est composé de K ratios économiques et financiers retenus pour leur qualité discriminante et leur faible corrélation entre elles. Le modèle requiert que les ratios choisis soient liés au risque de l'entreprise de façon linéaire. Le modèle estimé s'écrit :

$$y_i = 1 \text{ si } \beta x_i + \varepsilon_i > 0$$

$$y_i = 0 \text{ si } \beta x_i + \varepsilon_i \leq 0$$

où ε_i exprime l'erreur associée à l'entreprise i et β le vecteur des coefficients.

Des probabilités de défaut *a posteriori* sont alors aisément calculables ; la probabilité *a posteriori* que l'entreprise i soit défaillante est égale à :

$$\begin{aligned} P\{y_i = 0\} &= P\{\beta x_i + \varepsilon_i \leq 0\} = P\{(\varepsilon_i \leq -\beta x_i)\} \\ &= F(-\beta x_i) \end{aligned}$$

où F est la fonction de répartition des erreurs ε . Rappelons que la loi normale et la loi logistique sont symétriques. L'erreur ε est d'espérance nulle.

La probabilité *a posteriori* que l'entreprise i ne soit pas défaillante est :

$$\begin{aligned} P\{y_i = 1\} &= P\{\beta x_i + \varepsilon_i > 0\} = P\{\varepsilon_i > -\beta x_i\} \\ &= 1 - P\{\varepsilon_i \leq -\beta x_i\} = F(\beta x_i) \end{aligned}$$

Soit :

$$P\{y_i = 1\} = F(\beta x_i)$$

$$P\{y_i = 0\} = 1 - F(\beta x_i) = F(-\beta x_i)$$

Le modèle *Logit* suppose que les erreurs suivent une loi logistique ; leur fonction de répartition F s'écrit :

$$F(x) = (1 + e^{-x})^{-1}$$

La probabilité *a posteriori* que l'entreprise i soit défaillante s'écrit alors $P\{y_i = 0\} = (1 + e^{\beta x_i})^{-1}$

Le modèle *Probit* suppose que les erreurs suivent une loi normale ; leur fonction de répartition F s'écrit :

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

La probabilité *a posteriori* que l'entreprise i soit défaillante s'écrit alors $P\{y_i = 0\} = \int_{-\infty}^{-\beta x_i} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$

Comme précédemment, les probabilités *a posteriori* constituent des aides à la décision ; elles permettent la construction de classes de risque. Si l'on désire affecter une entreprise A à l'un des deux groupes à partir de la probabilité de faillite, il est possible de déterminer un seuil de décision, ici un seuil de probabilité P . Si sa probabilité de faillite estimée est supérieure à P , alors l'entreprise A est considérée comme défaillante ; inversement, si sa probabilité de faillite estimée est inférieure à P , elle est considérée comme non-défaillante. P est choisi de manière à maximiser la qualité du classement.

Les prévisions élaborées à partir des modèles *Logit* et *Probit* sont de bonne qualité : à un horizon d'un an, Platt et Platt (1991) aboutissent à des estimations du taux réel de bons classements égales à 85% pour les entreprises défaillantes et à 88% pour les entreprises non-défaillantes ; Mossman *et alii* (1998) estiment leur taux de bons classements des entreprises défaillantes égal à 80% et celui des entreprises non-défaillantes égal à 70%. Si Bardos (1989), comparant la discrimination linéaire aux modèles *Probit* à partir des mêmes données, tend à montrer la supériorité de la première, *a priori* grâce à sa plus grande robustesse, des études plus récentes aboutissent à la conclusion inverse. Par exemple, Kira *et alii* (1997) et Lennox (1999) obtiennent de meilleurs taux de bons classements respectivement par un modèle *Logit* et *Probit* que par un modèle d'analyse discriminante linéaire.

Les modèles recourant aux techniques non paramétriques de classification statistique

Lorsque les ratios ne suivent pas de lois multinormales et qu'une loi logistique ou une loi normale ne semble pas adaptée pour décrire les erreurs, une solution est de recourir aux méthodes non-paramétriques, ne nécessitant aucune hypothèse relative à la distribution des variables. Deux catégories de techniques ont été utilisées : le partitionnement récursif et l'analyse discriminante fondée sur une estimation locale de la fonction de distribution des ratios comptables.

Les modèles de prévision de la faillite fondés sur la technique du partitionnement récursif

La première méthode de classification non-paramétrique appliquée à la prévision de la faillite est l'algorithme de partitionnement récursif (Frydman *et alii*, 1985). La méthode consiste à construire un arbre décisionnel. L'échantillon, constitué d'entreprises défaillantes et non-défaillantes, est décomposé de manière séquentielle : à chaque nœud de l'arbre, un sous-groupe est décomposé en deux sous-groupes plus petits, jusqu'aux groupes finals, qui contiennent 100% d'entreprises défaillantes ou 100% d'entreprises non-défaillantes. On obtient alors des nœuds dits "purs". Un seuil de tolérance inférieur à 100% peut être accepté afin de limiter le nombre de nœuds, les nœuds finals sont alors "impurs". L'impureté d'un nœud rend compte de l'hétérogénéité du sous-groupe qui le caractérise.

À chaque nœud, le sous-groupe est scindé en deux sous-groupes au regard d'un ratio unique, qui doit avoir un lien monotone avec le risque de faillite. Un ratio et un seuil de décision sont choisis : les entreprises qui présentent une valeur de ce ratio inférieure au seuil sont classées dans le premier sous-groupe, les autres dans le second sous-groupe. Le ratio et le seuil sont fixés de manière à réduire au mieux l'impureté. L'un des sous-groupes contient

une proportion plus forte d'entreprises défaillantes et l'autre une proportion plus forte d'entreprises non-défaillantes. Chaque sous-groupe ainsi formé est ensuite également décomposé en deux autres sous-groupes, à l'aide d'un autre ratio, jusqu'à ce que la pureté du nœud – donc l'homogénéité des sous-groupes d'entreprises – soit jugée suffisante. Afin de classer une entreprise quelconque, il suffit de lui appliquer l'arbre décisionnel. Elle est considérée défaillante si *in fine* elle se situe dans un groupe constitué majoritairement d'entreprises défaillantes. Soulignons que cette méthode permet de prendre en compte le coût des erreurs de classement et les probabilités de défaut *a priori*.

Pour un horizon temporel de un an, Frydman *et alii* (1985) aboutissent à un taux de bons classements égal à 89% ou 80%⁽²⁷⁾. Bardos (1989) obtient un taux de bons classements égal à 69,6% pour les entreprises non-défaillantes ; pour les entreprises défaillantes, le taux est égal à 74,9% à un an de la faillite et égal à 61,2% à trois ans de la faillite. Même si dans l'étude de Bardos (1989) le partitionnement récursif est un peu moins efficace que l'analyse discriminante linéaire réalisée sur les mêmes données, cette méthode de classification est performante.

Le partitionnement récursif est cependant peu utilisé dans la prévision de la faillite. Une des explications envisageables est que cette méthodologie, contrairement à l'analyse discriminante linéaire, est très sensible au choix des probabilités de défaut *a priori* et des coûts d'erreur de classement. Ces variables conditionnent en effet à chaque nœud le ratio sélectionné. De plus, le partitionnement récursif requiert qu'à chaque nœud soit déterminée une valeur seuil pour le ratio considéré, ce qui nuit à la stabilité temporelle de l'indicateur.

Les modèles de prévision de la faillite fondés sur une estimation non-paramétrique des lois de distribution des ratios

Si l'application de la règle de décision bayésienne à l'analyse discriminante nécessite que soient connues les fonctions de densité des ratios, elle n'exige cependant pas que ces fonctions soient multinormales. Afin de résoudre le problème de la non-multinormalité des ratios, une autre loi de distribution est envisageable. Il est notamment possible d'estimer la loi de probabilité des ratios par des méthodes non-paramétriques⁽²⁸⁾ (Murphy et Moran, 1986, puis Calia et Ganuci, 1997).

Rappelons que les méthodes paramétriques d'estimation des lois de probabilités consistent à considérer une loi donnée, par exemple la loi normale, puis à évaluer empiriquement ses paramètres, en l'occurrence l'espérance et la variance ; elles supposent qu'un nombre fini de paramètres parviennent à décrire de façon suffisamment fidèle l'intégralité des évolutions

possibles de la variable considérée. Le principe des méthodes non-paramétriques est, au contraire, de ne pas s'appuyer sur une forme spécifique de loi déterminée *a priori*. Elles supposent qu'en présence de points extrêmes, les méthodes paramétriques en rendent compte imparfaitement. L'idée est alors de lisser les observations afin de mettre en évidence les évolutions principales. Le lissage consiste à estimer localement la fonction de densité par une moyenne pondérée locale des fréquences. Un point est pondéré d'autant plus fortement qu'il est proche du point x en lequel la fonction de densité est estimée. Le paramètre de lissage est crucial car il détermine la taille du voisinage, c'est-à-dire le nuage de points pris en considération pour estimer localement la fonction de densité ; il fixe également la pondération de chaque point. Il conditionne donc la précision de l'estimation.

Calia et Ganuci (1997) utilisent la méthode du noyau et la méthode du plus proche voisin, qui sont les deux modalités d'estimation les plus utilisées. Elles se différencient par la détermination du voisinage. La méthode du noyau considère un intervalle de largeur donnée autour du point x , le nombre de points considérés pour l'estimation étant variable. La méthode du plus proche voisin, au contraire, ne prend en considération que les k points les plus proches de x , l'intervalle étant variable. Après avoir ainsi estimé la fonction de densité de chaque ratio pour l'échantillon des entreprises défaillantes et des entreprises non-défaillantes, les auteurs appliquent la règle de décision de Bayes afin de classer les entreprises. Les paramètres de lissage de chaque méthode sont fixés de manière à obtenir la meilleure prévision possible.

À un an de la faillite, le taux global de bons classements estimé par Calia et Ganuci (1997) à partir d'un échantillon test distinct de l'échantillon initial est égal à 68% lorsque l'estimateur du noyau est utilisé et il est égal à 74,2% pour l'estimateur du plus proche voisin. L'indicateur ainsi élaboré est plus performant que l'analyse discriminante linéaire et quadratique sur les mêmes données, puisque ces techniques fournissent un taux de bons classements, estimé dans les mêmes conditions, respectivement égal à 60,9% et 61,8%. Le principal défaut de cette méthode statistique est le nombre important de données qu'elle requiert, par rapport à l'estimation paramétrique de la loi de distribution des ratios.

Les algorithmes de l'intelligence artificielle

Face aux nombreuses contraintes liées aux méthodes statistiques traditionnelles, des méthodologies relevant d'une logique différente ont été utilisées : les réseaux de neurones et les algorithmes génétiques. Elles relèvent de l'intelligence artificielle, plus précisément de la branche relative à l'apprentissage automatique. La première application des réseaux de neurones à l'estimation du risque de faillite a été réalisée sur données

bancaires (Bell *et alii*, 1990). L'utilisation des réseaux de neurones puis des algorithmes génétiques à partir de données d'entreprises non-financières s'est ensuite intensifiée.

La prévision de la faillite fondée sur les réseaux de neurones

Le principe des réseaux de neurones consiste en l'élaboration d'un algorithme dit d'apprentissage qui imite le traitement de l'information par le système neurologique humain⁽²⁹⁾. Chaque neurone réside en une fonction, dite de transfert, qui traite un ensemble d'informations (les *inputs*) afin d'obtenir un résultat (l'*output*). La fonction de transfert, choisie par l'utilisateur, consiste en un traitement mathématique, généralement non-linéaire, de l'*input*. À chaque *input* est attribué un poids qui influence le résultat. L'objectif est d'aboutir, après une phase dite d'apprentissage, à la combinaison des poids d'*inputs* de chaque neurone qui conduit à la meilleure description de la réalité, c'est-à-dire au meilleur classement des entreprises dans les deux groupes de firmes, les défaillantes et les non-défaillantes.

Trois sortes de neurones existent : les neurones d'entrée, les neurones de sortie et les neurones cachés. Les neurones d'entrée ont pour *inputs* les K ratios comptables pré-sélectionnés ; les neurones de sortie ont pour *output* la variable dichotomique défaillant / non-défaillant. Les neurones cachés sont des neurones qui traitent l'information entre les neurones d'entrée et les neurones de sortie. Le réseau utilisé peut être plus ou moins complexe, c'est-à-dire être constitué d'un nombre variable de couches de neurones. Une couche est un ensemble de neurones qui ne s'échangent pas d'information entre eux. En revanche, au sein de chaque couche les neurones sont connectés aux neurones de la couche précédente : l'ensemble des *outputs* d'une couche constitue par conséquent l'ensemble des *inputs* de la couche suivante. Les poids attribués à chaque *input* déterminent ainsi la transmission de l'information. L'analyse discriminante peut être considérée comme un réseau de neurones constitué d'un neurone unique dont les *inputs* seraient les valeurs des ratios comptables et l'*output* la valeur du score, *output* généré par transformation linéaire, qui constitue la fonction de transfert.

Pendant la phase d'apprentissage, le réseau est appliqué successivement à toutes les entreprises de l'échantillon. L'erreur de classement est calculée puis les poids des *inputs* sont modifiés de manière à réduire cette erreur ; c'est en effet par la variation progressive de ces pondérations que se réalise le processus d'apprentissage. L'ensemble de ces étapes est réitéré jusqu'à l'obtention de la classification optimale, c'est-à-dire de la minimisation de l'erreur de classement. À la fin de la phase d'apprentissage, le réseau permet de classer une entreprise quelconque dans le groupe des entreprises

défaillantes ou dans celui des entreprises non-défaillantes, avec un seuil d'erreur connu.

Les différents modèles de prévision de la faillite qui recourent aux réseaux de neurones se distinguent par la complexité du réseau (c'est-à-dire son nombre de couches et le nombre de neurones par couche) et par la fonction de transfert utilisée⁽³⁰⁾. De nombreuses études récentes se sont attachées à comparer les performances des réseaux de neurones aux techniques statistiques citées précédemment. Certaines tendent à montrer leur supériorité par rapport à l'analyse discriminante (Chung et Tam, 1993 ; Coats et Fant, 1993). L'étude d'Altman *et alii* (1994) présente un bilan plus mitigé : à un an de la faillite, les réseaux de neurones aboutissent à un taux de bons classements égal à 91,8% pour les entreprises défaillantes et à un taux égal à 95,3% pour les entreprises non-défaillantes. L'analyse discriminante linéaire menée sur le même échantillon aboutit à un classement légèrement meilleur : 92,8% pour les entreprises défaillantes et 96,5% pour les entreprises non-défaillantes. Kira *et alii* (1997) aboutissent à la même conclusion. Le taux global de bons classements à un an de la défaillance est égal à 93,3% lorsque l'analyse discriminante est utilisée ; il est égal à 90% avec un réseau de neurones. La régression *Logit* est plus performante, fournissant un taux de bons classements égal à 95,5%. Cependant, les différences sont ténues, comme le confirme l'étude de Bardos et Zhu (1997) : l'analyse discriminante linéaire classe correctement 68,3% des entreprises défaillantes et 75,3% des entreprises non-défaillantes alors que le réseau de neurones le plus performant classe correctement 61,3% des entreprises défaillantes et 82,8% des entreprises non-défaillantes.

Les réseaux de neurones permettent donc une prévision efficace de la faillite des entreprises. L'avantage principal de ce procédé consiste en l'absence de restrictions statistiques relatives à la distribution des variables et des erreurs et en l'absence de spécification *ex ante* d'une forme fonctionnelle. Cependant, plusieurs problèmes existent : le mécanisme d'apprentissage peut être long ; la robustesse du réseau à des modifications de l'environnement économique n'est pas encore prouvée, une mise à jour fréquente peut s'avérer nécessaire ; enfin la logique qui sous-tend le réseau est difficile à interpréter alors que la fonction score fournit des outils simples d'analyse de la défaillance.

La prévision de la faillite fondée sur les algorithmes génétiques

Varetto (1998) et Barney *et alii* (1999) ont les premiers envisagé une autre méthode, relevant également de l'intelligence artificielle : les algorithmes génétiques⁽³¹⁾. Le fonctionnement est assez proche de celui des réseaux de neurones. Un algorithme traite les informations disponibles, les

ratios comptables sélectionnés et la situation de chaque entreprise (défaillante ou non défaillante). L'algorithme, par itérations successives, déduit des régularités, des règles générales du processus de défaillance qu'il peut ensuite appliquer à d'autres entreprises⁽³²⁾.

Le principe d'un algorithme génétique est d'imiter le processus d'évolution naturelle des espèces. Il se fonde sur les trois principes du Darwinisme : sélections et reproductions des individus les "meilleurs", croisements génétiques et mutations aléatoires des gènes. L'algorithme génétique analyse donc des populations de solutions – c'est-à-dire des modes de traitement de l'information propres à prévoir la défaillance – et évalue leur qualité, en l'occurrence le taux de bons classements des entreprises qu'elles fournissent. L'algorithme conserve alors les modes de résolution les plus efficaces et les combine, éventuellement en génère de nouveaux de façon aléatoire. Il aboutit alors à une nouvelle population de solutions, plus homogène. La procédure est réitérée et elle s'arrête lorsque la population de solutions atteint un certain niveau d'homogénéité, fixé préalablement, tel qu'il est difficile d'améliorer "l'espèce" par de nouveaux croisements. L'algorithme peut finalement classer toute entreprise dans le groupe des entreprises défaillantes ou dans le groupe des entreprises non défaillantes.

Varetto (1998) compare l'efficacité d'un algorithme génétique à celle d'une analyse discriminante linéaire. Il considère deux formes d'algorithmes génétiques. Le premier algorithme génère des combinaisons linéaires de ratios comptables. Les coefficients des ratios ainsi que la constante sont successivement modifiés afin d'aboutir à la combinaison la plus discriminante. Le second algorithme génère une fonction score non-linéaire. Les résultats sont presque aussi bons que ceux fournis par l'analyse discriminante linéaire alors que la méthodologie de l'algorithme génétique n'exige pas les restrictions statistiques induites par l'analyse discriminante. Barney *et alii* (1999) incorpore un algorithme génétique dans un réseau de neurones. Les taux de bons classements obtenus sont supérieurs à ceux donnés par une régression *Logit*, effectuée sur le même échantillon. Shin et Lee (2002) obtiennent également des résultats prometteurs – puisque cette technique est encore dans une phase exploratoire. Cependant, les algorithmes génétiques présentent le même inconvénient que les réseaux de neurones, la difficile interprétation du rôle joué par chaque variable dans le processus de défaillance. Soulignons néanmoins que l'article de Shin et Lee (2002) prouve que les algorithmes génétiques peuvent fournir des règles de décision claires.

Conclusion

Les recherches académiques de méthodes de prévision de la faillite restent actives, pendant que leur utilisation opérationnelle s'intensifie. Nous pourrions rappeler les avantages – systématisation du traitement de l'information, gain de temps et de coût – et les inconvénients – problème de la robustesse temporelle et sectorielle, problèmes statistiques – communs à l'ensemble des techniques évoquées⁽³³⁾, mais notre objet est plutôt de mettre en perspective les différentes méthodes.

L'analyse discriminante linéaire, après une phase exploratoire lors des années soixante-dix, est la méthode la plus utilisée du point de vue opérationnel. Elle fournit en effet pour le moment les prévisions les plus robustes. De plus, la fonction score qu'elle permet d'établir est riche d'applications utiles pour les praticiens, comme le calcul de probabilité *a posteriori* ou la construction de classes de risques⁽³⁴⁾.

D'autres méthodologies ont été mises au point comme l'analyse discriminante quadratique, les régressions sur variables qualitatives ou encore les techniques non-paramétriques comme l'estimateur fondé sur la méthode du noyau. Leur objectif était principalement d'éviter les contraintes statistiques imposées par la discrimination linéaire. Ces techniques ne sont néanmoins pas aussi utilisés dans la pratique que la discrimination linéaire, principalement à cause de problèmes de robustesse (pour l'analyse discriminante quadratique) ou de disponibilité des données (pour les techniques non-paramétriques). Cependant, la qualité des prévisions fournies par les régressions sur variables qualitatives est proche de celle offerte par l'analyse discriminante linéaire.

Les récentes techniques empruntées à l'intelligence artificielle, tels les réseaux de neurones et les algorithmes génétiques, connaissent un très grand engouement académique. Elles permettent en effet de bonnes prévisions tout en présentant l'avantage de ne pas exiger de restrictions statistiques. De plus, les algorithmes génétiques sont particulièrement robustes, notamment parce qu'ils ne sont pas soumis à des contraintes mathématiques telles que la dérivabilité ou la continuité des fonctions utilisées⁽³⁵⁾. Leur utilisation, principalement celle des algorithmes génétiques, reste exploratoire. Leur haut degré de technicité ne leur permettait pas de fournir d'éléments explicatifs clairs de la défaillance, inconvénient qui semble pouvoir être réduit (Shin et Lee, 2002).

Une critique souvent formulée (Dumontier, 1991, par exemple) porte sur l'utilisation exclusive de données comptables : une amélioration significative de la détection de la faillite serait sans doute plus apportée par l'exploitation d'informations autres

que comptables et financières que par une complexification croissante des techniques de prévision. Des travaux récents amendent dans ce sens les méthodes présentées dans cet article. Citons par exemple l'étude de Gudmundsson (1999) qui améliore la qualité de sa prévision en ajoutant aux variables quantitatives "traditionnelles" des variables qualitatives, obtenues par enquête, relatives à la stratégie globale des entreprises. Le développement des systèmes-experts s'inscrit dans la même logique. Ils sont construits par l'application des arbres de décision puis des réseaux neuronaux et des algorithmes génétiques (voir Kim et Han, 2003, par exemple) à des critères qualitatifs utilisés par les experts, tels que la qualité de la gestion des ressources humaines ou le degré de concentration des clients. Les systèmes-experts permettent non seulement une compréhension plus riche du processus de défaillance mais également une bonne prévision de la faillite. Rappelons néanmoins pour finir que les modèles de prédiction du risque de faillite ne constituent qu'une aide à la décision parmi d'autres et se doivent d'être enrichis par d'autres formes d'information.

Notes

(1) Cf. par exemple Guilhot (2000) pour une présentation globale des analyses relatives à la faillite.

(2) En France, la procédure judiciaire actuelle s'appelle une procédure de redressement judiciaire, qu'elle aboutisse à une liquidation, une cession ou un plan de redressement. Aux États-Unis, une entreprise peut déposer son bilan dans le cadre du *Chapter 11* du code des procédures collectives si seul un plan de restructuration semble nécessaire ou dans le cadre du *Chapter 7* si une liquidation semble inévitable.

(3) Un *covenant* est une clause d'un contrat de prêt limitant l'action du dirigeant dans l'intérêt du créancier. Elle peut porter sur la vente d'actifs, le versement de dividendes, l'émission ultérieure de dette de même rang, les investissements mis en place etc.

(4) Cf. tableau récapitulatif.

(5) Altman *et alii* (1994) ont conjugué les deux dernières approches en réalisant une double classification : la première distingue les entreprises saines des entreprises fragiles, puis la seconde détecte, au sein des entreprises fragiles, celles qui sont susceptibles de faire l'objet d'une procédure judiciaire.

(6) Ceci est vrai lorsqu'un échantillon rétrospectif – *i.e.* tel que chaque sous-échantillon est obtenu par tirage aléatoire au sein de chaque sous-population – est utilisé et que le taux de sondage ne respecte pas le taux de défaillance de la population globale (cf. Bardos et Zhu, 1997).

(7) Cette méthode est souvent plus robuste. En effet, lorsque les hypothèses de multinormalité et d'homoscédasticité des échantillons ne sont pas respectées, les résultats de la classification sont moins biaisés lorsque les deux sous-échantillons sont de tailles similaires.

(8) Le recours aux données relatives aux cours boursiers est difficile en France, où un pourcentage très faible des entreprises se finance sur les marchés.

(9) Le capital engagé est égal à la somme de l'endettement financier total, du financement propre et des dettes du groupe et des associés.

(10) Le capital économique est égal à la somme des immobilisations productives et du besoin en fonds de roulement d'exploitation.

(11) Le *cash-flow* de l'entreprise peut être défini comme le flux de trésorerie global, égal à l'excédent brut global réduit de la variation des besoins en fonds de roulement (Charreaux, 2000).

(12) L'intégralité des ratios que nous venons de citer, qu'ils rendent compte de la rentabilité, de la structure ou de la capacité de remboursement, présente l'avantage d'avoir un dénominateur systématiquement positif. Le problème d'un numérateur et d'un dénominateur simultanément négatifs ne se pose donc pas.

(13) Le risque spécifique est la part de la variance du cours d'une action qui n'est pas liée aux évolutions du marché des titres dans son ensemble mais à des facteurs propres à l'entreprise.

(14) Le nombre de banques est fortement lié à la taille de l'entreprise. Les auteurs ont donc constitué deux sous-échantillons : les entreprises de plus de 50 employés et les entreprises de moins de 50 employés. Le nombre de banques ne permet d'améliorer la prévision que pour ce dernier échantillon.

(15) Pour plus de détails, voir par exemple Efron et Tibshirani (1993).

(16) La prédiction de la défaillance des banques s'est vue appliquer d'autres méthodes : la technique d'enveloppement des données (Barr *et alii*, 1994) et les modèles de durée de vie (Henebry, 1996), qui permettent de prévoir l'échéance de la faillite. Cette dernière méthode a été appliquée par Li (1999) afin d'estimer la durée d'un redressement des entreprises industrielles soumises au *Chapter 11* aux États-Unis et par Honjo (2000) afin de déterminer les causes de faillite des entreprises jeunes.

(17) Pour plus de détails, voir par exemple Bardos (2001).

(18) Afin de tenir compte de la covariance entre les variables, la distance de Mahalanobis est souvent utilisée. Soit Ω la matrice de variance-covariance des K variables explicatives considérées. La distance de Mahalanobis entre l'entreprise A et le groupe j est alors $(x_A - x_{\text{ker}}^j)' \Omega (x_A - x_{\text{ker}}^j)$ où x_{ker}^j est le vecteur des coordonnées du centre de gravité du groupe j .

(19) Une liste exhaustive des articles est difficile à établir et surtout trop longue pour être présentée ici. Nous ne citons que les études qui nous semblent les plus importantes en termes d'amélioration de la prédiction de la faillite : Lev (1969) ; Eisenbeis et Avery (1972) ; Trieschmann et Pinches (1973) ; Altman et Loris (1976) ; Altman *et alii* (1977) ; Taffler (1982) ; Izan (1984) ; Rose et Giroux (1984) ; Gentry *et alii* (1985) ; Platt et Platt (1990 et 1991) ; Reuilly (1990) ; Bardos (1998a).

(20) Pour un recensement international – non exhaustif – des modèles de classification des entreprises selon leur risque de faillite, voir Altman *et alii* (1997).

(21) Le problème principal de la prise en compte des coûts des erreurs de classification consiste en leur quantification. Altman *et alii* (1977) – première étude qui les intègre – estime que le rapport de C_D^{ND} à C_D^{ND} est égal à 35, Malécot (1986) l'estime égal à 30. Taffler (1982) utilise un rapport entre le coût des erreurs de type I et le coût des erreurs de type II de l'ordre de 200.

(22) Rappelons que la probabilité de défaut *a priori* de l'entreprise A est la probabilité estimée sans qu'aucune information ne soit connue ; elle est égale au taux de défaillance observé sur la population totale des entreprises.

(23) Voir Bardos (2001).

(24) Cf. les études relatives aux caractéristiques des entreprises défaillantes, comme par exemple Blazy *et alii* (1993) ou Augory *et alii* (1996).

(25) Cf. par exemple la critique de Malécot (1986).

(26) Voir par exemple Greene (2000) pour plus de détails.

(27) Le chiffre de 89% est obtenu en considérant un rapport entre le coût des erreurs de type I et celui des erreurs de type II égal à 20. Le taux de bon classement est plus faible (80%) lorsque ce rapport est égal à 50.

(28) Pour plus de détails, voir par exemple Yatchew (1998).

(29) Pour une description détaillée des principaux types de réseaux neuronaux et de fonctions de transfert utilisés, voir par exemple Cottrell *et alii* (1996) ou Blayo et Verleysen (1996).

(30) À titre d'illustration, Altman *et alii* (1994) utilisent deux réseaux de neurones pour reconstituer à partir de ratios comptables les valeurs du score élaboré par analyse discriminante linéaire et pour discriminer les entreprises défaillantes des entreprises non défaillantes. Le premier réseau est constitué de trois couches : une de dix neurones, une de quatre et la dernière, destinée à produire l'*output* final, constituée d'un neurone unique. 1000 cycles d'itérations ont été nécessaires au processus d'apprentissage. Le second réseau est également constitué de trois couches, contenant respectivement 15, 6 et 1 neurones ; 2000 cycles de traitement ont été nécessaires.

- (31) Voir par exemple Goldberg (1994).
 (32) Pour une illustration simple, voir l'application au choix d'un portefeuille d'actions proposée par Alexandre (2001).
 (33) Voir notamment Malécot (1986), Dumontier (1991) ou Grice et Dugan (2001).
 (34) Voir Bardos (2001).
 (35) Voir Alexandre (2001).

Bibliographie

- Aharony J., Jones C. et Swary I. (1980).** "An analysis of risk characteristics of corporate bankruptcy firms using capital market data", *Journal of Finance*, vol. 35, pp. 1001-1016.
- Alexandre H. (2001).** *L'individu face au marché. Investisseurs, spéculateurs et crises boursières*, Éditions EMS, collection "L'actualité de la Gestion".
- Altman E.I. (1968).** "Financial Ratios, Discriminant Analysis and the Predictions of Corporate Bankruptcy", *Journal of Finance*, vol. 23, n° 4, pp. 589-609.
- Altman E.I. (1977).** "Predicting Performances in the Saving and Loans Association Industry", *Journal of Monetary Economics*, vol. 3, pp. 433-466.
- Altman E.I. (2002).** "Corporate distress prediction models in a turbulent economic and Basel II environment", *Document de travail*.
- Altman E.I., Haldeman R.G. et Narayan P. (1977).** "Zeta analysis: A New Model to Identify Bankruptcy Risk of Corporations", *Journal of Banking and Finance*, vol. 1, Juin, pp. 29-54.
- Altman E.I. et Brenner M. (1981).** "Information effects and stocks market response to signs of firm deterioration", *Journal of Financial and Quantitative Analysis*, vol. 16, pp. 35-51.
- Altman E.I. et Loris B. (1976).** "A financial early warning system for the over-the-counter broker-dealers", *Journal of Finance*, vol. 31, pp. 1201-1217.
- Altman E.I., Marco G. et Varetto F. (1994).** "Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience)", *Journal of Banking and Finance*, vol. 18, pp. 505-529.
- Altman E.I. et Narayan P. (1997).** "An international survey of business failure classification models", *Financial Markets, Institutions and Instruments*, vol. 6, n° 2, pp. 1-54.
- Augory C., Avouyi-Dovi S., Busque P. et Quéron A. (1996).** "Les défaillances d'entreprises au début des années 90" *Lettre économique de la CDC*, Octobre.
- Aziz A., Emmanuel D. et Lawson G. (1988).** "Bankruptcy prediction – an investigation of cash flow based models", *Journal of Management Studies*, vol. 25, pp. 55-63.
- Bardos M. (1989).** "Trois méthodes d'analyse discriminante", *Cahiers Économiques et Monétaires*, Banque de France, 2^{ème} trimestre, n°33.
- Bardos M. (1998a).** "Detecting the risk of company failure at the Banque de France", *Journal of Banking and Finance*, vol. 22, pp. 1405-1419.
- Bardos M. (1998b).** "Le Score BDFI – Du diagnostic individuel à l'analyse de portefeuille", *Les Études de l'Observatoire des Entreprises*, Banque de France, janvier, 76 pages.
- Bardos M. (2001).** *Analyse discriminante : Application au rating et scoring financier*, Dunod.
- Bardos M. et Zhu W.H. (1997).** "Comparaison de l'analyse discriminante linéaire et des réseaux de neurones – Application à la détection de défaillance d'entreprises", *Revue Statistique Appliquée*, vol. XLV, n° 4, pp. 65-92.
- Barney D.K., Graves O.F. et Johnson J.D. (1999).** "The farmer home administration and farm failure prediction", *Journal of Accounting and Public Policy*, vol. 18, pp. 99-139.
- Barr R.S., Seiford L.M. et Siems T.F. (1994).** "Forecasting Bank Failure: A non-Parametric Frontier Estimation Approach", *Recherches Économiques de Louvain*, vol. 60, n° 4, pp. 417-429.

- Beaver W. (1966).** "Financial Ratios as Predictors of Failure", *Empirical Research in Accounting: Selected Studies, Journal of Accounting Research*, supplément au vol. 4, pp. 71-111.
- Bell T., Ribar G. et Verchio J. (1990).** "Neural networks versus logistic regression in predicting bank failures", in P. Srivastava Ed., *Auditing Symposium X*, University of Kansas.
- Blayo F. et Verleysen M. (1996).** *Les réseaux de Neurones Artificiels*, Que sais-je ?
- Blazy R., Charletty P. et Combier J. (1993).** "Les défaillances d'entreprises : des difficultés visibles plusieurs années à l'avance", *INSEE Economie et Statistiques*, n° 268-269, pp. 101-109.
- Burgstahler D., Jiambalvo J. et Noreen E. (1989).** "Changes in the Probability of Bankruptcy and Equity Value", *Journal of Accounting and Economics*, vol. 11, pp. 207-224.
- Calia P. et Ganugi P. (1997).** "Kernel and nearest neighbour discriminant analysis: Business failure classification in industrial district", *Applied Stochastic Models and Data Analysis*, Colloque Capri 1997.
- Charreaux G. (2000).** *Gestion Financière*, Litec, Collection "Expertise Comptable".
- Chung H. et Tam K.Y. (1993).** "A comparative analysis of inductive learning algorithms", *Intelligence Systems in Accounting, Finance and Management*.
- Clark T. et Weinstein M. (1983).** "The behaviour of common stock of bankrupt firms", *Journal of Finance*, 38, pp. 489-504.
- Coats P. et Fant L. (1993).** "Recognizing financial distress patterns using a neural network tool", *Financial Management*, novembre, pp. 142-155.
- Collins R. (1980).** "An empirical comparison of bankruptcy prediction models", *Financial Management*, Eté, pp. 51-57.
- Cottrell M., de Bodt E. et Levasseur M. (1996).** "Réseaux de neurones en finance : Principes et revue de la littérature", *Finance*, vol. 16, pp.25-92.
- Deakin E.B. (1972).** "A Discriminant Analysis of Predictors of Business Failure", *Journal of Accounting Research*, vol. 10, n°1, printemps, pp. 167-179.
- Dumontier P. (1991).** "Vices et vertus des modèles de prévision de défaillance", *La Revue Banque*, n°514, mars, pp. 275-282.
- Efron B. et Tibshirani R.J. (1993).** *An Introduction to the bootstrap*, Chapman & Hall Ed.
- Eisenbeis R. et Avery R.B. (1972).** "Discriminant Analysis and Classification Procedures: Theory and Applications", *Lexington*.
- Flagg J.C., Giroux G.A. et Wiggins C.E. Jr.(1991).** "Predicting Corporate Bankruptcy using failing firms", *Review of Financial Studies*, vol. 1, n°1, pp. 67-78.
- Foglia A., Laviola S., Marullo Reedtz P. (1998).** "Multiple banking relationships and the fragility of corporate borrowers", *Journal of Banking and Finance*, vol. 22, n° 10-11, pp. 1441-1456.
- Frydman H., Altman E. et Kao D.-L. (1985).** "Introducing recursive partitioning for financial classification: the case of financial distress", *Journal of Finance*, vol.40, n°1, pp. 269-291.
- Gebhardt G. (1980).** "Insolvency Prediction Based on Annual Financial Statements According to the Company Law – An Assessment of the Reform of Annual Statements by the Law of 1956 from the View of External Addresses", in H. Bester et alii, eds. *Bochumer Beitrage Zur Untennehmungs und Unternehmens-forschung*, vol. 22.
- Gentry J.A., Newblod P. Et Whiteford D.T. (1985).** "Classifying Bankrupt Firms With Funds Flow Components", *Journal of Accounting Research*, vol. 23, n°1, pp. 146-160.
- Goldbger D.E. (1994).** *Algorithmes génétiques. Exploration, optimisation et apprentissage*, Ed. Addison-Wesley.
- Greene W.H. (2000).** *Econometrics Analysis*, Prentice Hall, 4^{ème} édition.
- Grice J.S. et Dugan M. T. (2001).** "The limitations of bankruptcy models: some cautions for the researcher", *Review of Quantitative Finance and Accounting*, 17, pp. 151-166.
- Gudmundsson S. V. (1999).** "Airline failure and distress prediction: a comparison of quantitative end qualitative models", *Transportation Research Part E*, 35, pp. 155-182.
- Guilhot B. (2000).** "Défaillances d'entreprises – Soixante-dix ans d'analyses théoriques et empiriques", *Revue Française de Gestion*, n° 130, septembre-octobre, pp. 52-67.
- Henebry K.-L. (1996).** "Do Cash Flow Variables Improve the Predictive Accuracy of a Cox Proportional Hazards Model for Bank Failure?", *Quarterly Review of Economics and Finance*, vol. 36, n° 3, pp. 395-409.
- Honjo Y. (2000).** "Business failure of new firms: An empirical analysis using a multiplicative hazards model", *International Journal of Industrial Organization*, vol. 18, pp. 557-574.
- Houghton K.A. (1984).** "Accounting Data and the Prediction of Business Failure: The Setting of Priors and the Age of Data", *Journal of Accounting Research*, vol. 22, n° 1, printemps, pp. 361-368.
- Izan H. (1984).** "Corporate distress in Australia", *Journal of Banking and Finance*, vol. 8, pp. 303-320.
- Kim M.-J. et Han I. (2003).** "The discovery of experts' decision rules from qualitative bankruptcy data using genetic algorithms", *Expert Systems with Applications*, 25, pp. 637-646.
- Kira D.S., Doreen D. et Nguyen D. (1997).** "An application of artificial neural networks and statistical methods in qualitative evaluation of small business loans", *Applied Stochastic Models and Data Analysis*, Colloque Capri 1997.
- Lachenbruch P.A., Sneider C.A. et Revo L.T.(1973).** "Robustness of the linear and quadratic discriminant function to certain type of non-normality", *Commun.Statistique.*, vol. 1, pp. 39-56.
- Lachenbruch P.A. (1975).** *Discriminant Analysis*, Hafner Press, New York.
- Lennox C. (1999).** "Identifying Failing Companies: A Revaluation of the Logit, Probit and DA approaches", *Journal of Economics and Business*, vol. 51, pp. 347-364.
- Lev B. (1969).** "Industry averages as targets for financial ratios", *Journal of Accounting Research*, vol. 7, pp. 31-51.
- Li K. (1999).** "Bayesian analysis of duration models: An application to Chapter 11 bankruptcy", *Economics Letters*, vol. 63, n° 3, pp. 305-312.
- Malécot J.-F. (1986).** "Sait-on vraiment prévoir les défaillances d'entreprises?", *Économies et Sociétés*, vol. 20, n° 12, décembre, pp. 55-82.
- Marks S. et Dunn O.J. (1974).** "Discriminant functions when covariance matrices are unequal", *Journal of American Statistic Association*, vol. 69, pp. 555-559.
- Mensah Y.M. (1984).** "An Examination of the Stationarity of Multivariate Bankruptcy Prediction Models: A methodological Study", *Journal of Accounting Research*, vol. 22, n°1, printemps, pp. 380-395.
- Michalopoulos M., Zopounidis C. et Kouris I. (1993).** "An integrated DSS for the assessment of the risk of failure", *Applied Stochastic Models and Data Analysis*, J. Janssen et C.H. Sjiadas Ed.
- Mossman C.E., Bell G.G. Swartz L.M. etTurtle H. (1998).** "An empirical comparison of bankruptcy models", *Financial Review*, vol. 33, pp. 35-54.

Murphy B.J. et Moran M.A. (1986). "Parametric and kernel density estimation in discriminant analysis: another comparison", *Comput. Math. Applic.*, 12A.

Ohlson J. (1980). "Financial Ratios and Ratios and the Probabilistic Prediction of Bankruptcy", *Journal of Accounting Research*, vol. 18, Printemps, pp. 109-131.

Platt H.D. et Platt M.B. (1990). "Improving ex ante bankruptcy classification using stable predictive variables", *Journal of Business Finance and Accounting*, vol. 17, pp. 31-51.

Platt H. D. et Platt M.B. (1991). "A note on the use of industry-relative ratios in bankruptcy prediction", *Journal of Banking and Finance*, vol. 15, pp. 1183-1194.

Reilly F.K. (1990). "Using Cash Flows and Financial Ratios to Predict Bankruptcies", in *Analysing Investment Opportunities in Distressed and Bankrupted Companies*, T.A. Bowman Ed., Charlottesville, VA.

Rose P. S. et Giroux G.A. (1984). "Predicting Corporate Bankruptcy: An Analytical and Empirical Evaluation", *Review of Business and Economic Research*, vol. XIX, n° 2, pp. 1-12.

Saporta G. (1977). "Une méthode et un programme d'analyse discriminante pas à pas sur variables qualitatives", *INRIA*, septembre.

Shin K.-S. Et Lee Y.J. (2002). "A genetic algorithm application in bankruptcy prediction modelling", *System Experts with Applications*, 23, pp. 637-646.

Taffler R.J. (1982). "Forecasting Company Failure in the UK using Discriminant Analysis and Financial Ratio Data", *Journal of Royal Statistical Society, Series A*, vol. 145, Part 3, pp. 342-358.

Tamari M. (1964). "Financial Ratios as a Means of Forecasting Bankruptcy", *Economic Review*, Bank of Israel, Jérusalem.

Trieschmann J.S. et Pinches G.E. (1973). "A multivariate model for predicting financially distressed P-L insurers", *Journal of Risk and Insurance*, vol. 40, pp. 327-338.

Varetto F. (1998). "Genetic Algorithms Applications in the Analysis of Insolvency Risk", *Journal of Banking and Finance*, vol. 22, n° 10-11, octobre, pp. 1421-39.

Weiss, L.-A. (1996). "The Impact of Incorporating the Cost of Errors Into Bankruptcy Prediction Models", *INSEAD Working Paper*, 96/27/AC, mai.

Yatchew A. (1998). "Nonparametric Regression Techniques in Economics", *Journal of Economic Literature*, vol. XXXVI, juin, pp. 669-721

Zmijewski M.E. (1984). "Methodological issues related to the estimation of financial distress prediction models", *Journal of Accounting Research*, Supplément au vol. 22, pp. 59-86.

Tableau récapitulatif des principales études citées

Le tableau ci-après propose un récapitulatif des principales études citées. Compte tenu de l'ampleur des travaux réalisés, à une liste exhaustive nous avons préféré une sélection des articles fondateurs des différentes méthodes, ainsi que les études les plus récentes. Pour chacune d'entre elles les caractéristiques de l'échantillon global sont décrites : la nature des entreprises (nationalité et secteur), la date des observations, la définition de la défaillance retenue et la structure de l'échantillon (c'est-à-dire la taille respective de la population des entreprises défaillantes et des entreprises non défaillantes). Ensuite sont présentés le nombre de ratios sélectionnés, la technique de classification retenue et les taux de bons classements. Ceux-ci sont donnés pour les entreprises défaillantes et pour les entreprises non défaillantes à un an de la défaillance et, entre parenthèses, à trois ans de la défaillance. La méthode de validation des résultats est citée afin de pouvoir comparer les études. Soulignons qu'une confrontation véritable des différentes méthodes de classification ne peut se faire que si elles ont été appliquées aux mêmes données. Nous avons donc privilégié dans notre sélection les études proposant une telle possibilité.

Définition de la défaillance		Structure de l'échantillon (D versus ND)	Nature et nombre de variables utilisées	Méthodologie utilisée	Méthode de validation	Taux de Bons Classements (#)
Incidents de paiement ou dépôt de bilan		79 D et 79 ND couplées par taille	14 ratios comptables	Classification unidimensionnelle	Echantillon test différent	Global 87 % (77%)
Dépôt de bilan		33 D et 33 ND couplées par taille et par secteur	5 ratios comptables	ADL	Echantillon initial	D 94% (n.d.) ND 97% (n.d.)
Dépôt de bilan		32 D et 32 ND	14 ratios comptables	Classification unidimensionnelle	Echantillon test différent	D 96% (n.d.) ND 79% (n.d.)
Dépôt de bilan		105 D et 2058 ND	9 ratios comptables	Logit	Echantillon initial	Global 87% (82%)
Dépôt de bilan		23 D et 45 ND	5 ratios comptables	ADL	<i>Lack-knife</i> (échantillon test de 33 D et 10 ND)	D 79,8% (n.d.) ND 91,4% (n.d.)
Élimination des ND vulnérables						D 87,9% (48%) ND 100% (n.d.)
Dépôt de bilan		46 D et 46 ND	18 ratios comptables	ADL	Echantillon test différent	D 84,6% (87,5%) ND 97,1% (96,2%)
Dépôt de bilan		couplées par secteur et par taille		ADQ		D 88,5% (93,8%) ND 80,0% (57,7%)
Dépôt de bilan		58 D et 142 ND sélectionnées aléatoirement	6 ratios comptables	Partitionnement récursif	<i>Bootstrap</i>	Global 85 % (n.d.)
Incidents de paiement ou signal défavorable		202 entreprises ayant connu des difficultés financières dont 26 ont déposé leur bilan	10 ratios comptables	ADL		Global 96,3% (n.d.)
			6 ratios comptables	Régression <i>Logit</i>	Echantillon initial	D 73% (n.d.) ND 97% (n.d.)
Dépôt de bilan		182 D et 182 ND couplées par taille et par secteur	7 ratios comptables relatifs	Régression <i>Logit</i>	Echantillon initial	D 85% (n.d.) ND 88% (n.d.)
						D 82% (n.d.) ND 78% (n.d.)
			7 ratios comptables absolus			
Difficultés de paiement ou procédure judiciaire		404 entreprises en redressement 404 entreprises en difficultés 404 entreprises saines	Ratios n.d.	ADL	Echantillon initial	D 92,8% (90,3%) ND 96,5% (86,4%) D 91,8% (89,4%) ND 95,3% (86,2%)
Ouverture d'une procédure judiciaire		809 D 1381 ND	Ratios comptables Nombre n.d.	ADL RN	Validation croisée	D 68,3% (n.d.) ND 75,3% (n.d.) D 61,3% (n.d.) ND 82,8% (n.d.)
Dépôt de bilan		41 D 917 ND	20 ratios comptables	ADL ADQ	Echantillon test différent	Global 60,9% (n.d.) Global 61,8% (n.d.)
Octroi de prêt ou non		36 entreprises à qui prêt refusé 99 entreprises à qui prêt accepté (Sélection aléatoire)	Ratios comptables et informations qualitatives Nombre n.d.	ADL <i>Logit</i> RN	Méthode du noyau Plus proche voisin Échantillon test différent	Global 68% (n.d.) Global 74,2% (n.d.) Global 93,3% (n.d.) Global 95,5% (n.d.) Global 90% (n.d.)

des ations	Définition de la défaillance	Structure de l'échantillon (D versus ND)	Nature et nombre de variables utilisées	Méthodologie utilisée	Méthode de validation	Taux de Bons Classements (*)
5	Dépôt de bilan	Environ 900 D et 38000 ND (échantillon non constant)	7 ratios comptables 5 Ratios comptables	Analyse discriminante	Échantillon test différent	D 78,1% (n.d.) ND 77,3% (n.d.) D 80,0% (48,7%) ND 71,1% (44,7%)
1	Dépôt de bilan	95 D et 95 ND couplées par taille et par secteur	5 Cash-flows Rendement des actions Variance du rendement des actions	Régression Logistique	Jack-Knife	D 71,1% (50%) ND 68,9% (63,2%) D 66,7% (60,5%) ND 57,8% (57,9%) D 62,2% (65,8%) ND 42,2% (50%)
5	Dépôt de bilan ou entreprise déclarée insolvable par une banque	1920 D et 1920 ND	Ratios comptables Nombre n.d.	ADL AG	Échantillon test différent	D 95,3% (n.d.) ND 93,5% (n.d.) D 96,5% (n.d.) ND 89,1% (n.d.)
2	Défaut de paiement	54 D et 184 ND	14 ratios comptables	Logit	Échantillon test différent	D 95% (n.d.) ND 85% (n.d.)
4	Dépôt de bilan	160 D et 789 ND	5 ratios comptables Variables de taille et de secteur	RN et AG Probit Logit ADL	Échantillon test différent	D 90% (n.d.) ND 95% (n.d.) D 72,7% (n.d.) ND 91,5% (n.d.) D 69,7% (n.d.) ND 93,3% (n.d.) D 78,8% (n.d.) ND 85,4% (n.d.)
7	Dépôt de bilan	264 D et 264 ND	9 ratios comptables	AG	Échantillon test différent	Global 80,8% (n.d.)

ance)